CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION CFSTI
DOCUMENT MANAGEMENT BRANCH 410.11

LIMITATIONS IN REPRODUCTION QUALITY

ACCESSION #

☒ 1. WE REGRET THAT LEGIBILITY OF THIS DOCUMENT IS IN PART UNSATISFACTORY. REPRODUCTION HAS BEEN MADE FROM BEST AVAILABLE COPY.

☐ 2. A PORTION OF THE ORIGINAL DOCUMENT CONTAINS FINE DETAIL WHICH MAY MAKE READING OF PHOTOCOPY DIFFICULT.

☐ 3. THE ORIGINAL DOCUMENT CONTAINS COLOR, BUT DISTRIBUTION COPIES ARE AVAILABLE IN BLACK-AND-WHITE REPRODUCTION ONLY.

☐ 4. THE INITIAL DISTRIBUTION COPIES CONTAIN COLOR WHICH WILL BE SHOWN IN BLACK-AND-WHITE WHEN IT IS NECESSARY TO REPRINT.

☐ 5. LIMITED SUPPLY ON HAND: WHEN EXHAUSTED, DOCUMENT WILL BE AVAILABLE IN MICROFICHE ONLY.

☐ 6. LIMITED SUPPLY ON HAND: WHEN EXHAUSTED DOCUMENT WILL NOT BE AVAILABLE.

☐ 7. DOCUMENT IS AVAILABLE IN MICROFICHE ONLY.

☐ 8. DOCUMENT AVAILABLE ON LOAN FROM CFSTI (TT DOCUMENTS ONLY).

☐ 9.

NBS 9 64 PROCESSOR:

604231

COMPUTATIONAL PROBLEMS IN THE THEORY
OF DYNAMIC PROGRAMMING

By

Richard Bellman

P-423

6 August 1953

The RAND Corporation
1700 MAIN ST. · SANTA MONICA · CALIFORNIA

Summary: A number of representative approximation techniques in the theory of dynamic programming are illustrated in a discussion of ~~the~~ equation

$$f(x) = \max_{0 \leq y \leq x} \left[g(y) + h(x-y) + f(ay + b(x-y)) \right].$$

COMPUTATIONAL PROBLEMS IN THE THEORY OF DYNAMIC PROGRAMMING

By

Richard Bellman

§1. Introduction.

In recent years, as multi-stage processes have come to assume a role of greater and greater importance in the industrial and economic arena, a number of interesting and novel mathematical problems have arisen, many of formidable caliber. The theory of dynamic programming was created to furnish an approach to these problems. The essential aim of the theory is to translate these questions from the unfamiliar field of policies, strategies, programming and scheduling, and such seeming imponderables, into functional equations which can be attacked by the precise techniques of analysis. These equations are, however, nonlinear in general, and possess the usual feature of problems which occur in applications, namely resolute and impartial insolubility.

Since a theory that has pretensions of application stands or falls upon its ability to produce numbers, it is of paramount importance to derive approximate techniques which may be used to determine numerical solutions.

In the following pages we shall consider a simple problem involving a sequence of decisions, first formulating it in classical form and then in terms of the dynamic programming approach. We shall then use this resultant functional equation to illustrate a number of approximation techniques, employing the particularly important concept of approximation in strategy space.

In closing we shall mention briefly some problems of more complicated form to which the same techniques are applicable.

§2. Optimal Allocation.

As a simple example of a large class of problems that occur in applications, let us consider the following. We are given a quantity $x > 0$ that may be divided into two parts y and $x-y$. From y we obtain a return of $g(y)$, and from $(x-y)$ a return of $h(x-y)$. In so doing we expend a certain amount of our original resources and are left with a new quantity $ay + b(x-y)$, where a and b are positive constants less than one, with which to continue the process. How does one proceed to maximize the total return obtained over N stages?

The conventional approach to this problem begins by listing the allocations $y_1, y_2, \cdots, y_N$ at the first, second, $\cdots$ and N^{th} stages. The total return from this sequence of choices will be

$$(1) \qquad J(y_1, y_2, \cdots, y_N) = \sum_{i=1}^{N} g(y_i) + \sum_{i=1}^{N} h(x_i - y_i),$$

where the variables are constrained by the conditions

$$(2) \qquad \begin{aligned} &0 \leq y_i \leq x_i, \\ &x_1 = x \\ &x_2 = ay_1 + b(x_1 - y_1) \\ &\vdots \\ &x_N = ay_{N-1} + b(x_{N-1} - y_{N-1}). \end{aligned}$$

The problem is now to maximize J subject to the above restrictions. Since several of the optimal y_i may be boundary points, and in some cases all <u>are</u> boundary points, an unrestricted use of calculus is not possible.

We are now confronted with a problem possessing the typical nasty features of maximization problems over N-dimensional regions. Furthermore, we observe that solving the problem in its present form yields too much information in the sense that we determine $y_1, y_2, \cdots,$ and y_N simultaneously, whereas all that is actually required to carry out the process is y_1 as a function of x and the number of stages remaining.

In the next section we shall formulate the problem from that point of view.

§3. Functional Equation Approach.

Let us define

(3) $f_N(x)$ = total return obtained from N-stages using an optimal policy

It is clear that the maximum over-all return is a function only of the initial amount x and the number of stages remaining.

If the initial allocation is y, the total return will be $g(y) + h(x-y)$ + the return from the succeeding (N-1) stages. Since it is easily seen that an optimal policy must have the property that its continuation after the first stage must be optimal with respect to the new initial amount $ay + b(x-y)$ and the remaining

(N-1) stages, we obtain as the total return due to an initial choice of y

$$(4) \qquad R_N(y) = g(y) + h(x-y) + f_{N-1}(ay + b(x-y)).$$

Since we wish to maximize the total return, y is now chosen to maximize this, yielding the functional equation

$$(5) \qquad f_N(x) = \operatorname*{Max}_{0 \leq y \leq x} R_N(y)$$

$$= \operatorname*{Max}_{0 \leq y \leq x} \left[g(y) + h(x-y) + f_{N-1}(ay + b(x-y)) \right],$$

for $N \geq 2$, with

$$(6) \qquad f_1(x) = \operatorname*{Max}_{0 \leq y \leq x} \left[g(y) + h(x-y) \right].$$

We shall assume henceforth that g and h are continuous functions in the interval $[0, \bar{x}]$, so that the maxima are all assumed.

We have thus replaced the original problem, as described in (1) and (2) of §2, by the sequence of recurrence relations in (5) and (6) above. Although these recurrence relations are nonlinear, the region of variation is one-dimensional. To justify this transformation of the original problem, we must show that these equations above can be utilized to yield both theoretical and numerical results.

§4. A Preliminary Approximation.

Let us begin by making a preliminary approximation that N is infinite. In place of the system of recurrence relations of (5) and (6) of §3 we obtain one functional equation

$$(7) \qquad f(x) = \operatorname*{Max}_{0 \leq y \leq x} \left[g(y) + h(x-y) + f(ay + b(x-y)) \right]$$

where

$$(8) \qquad f(x) = \lim_{N \to \infty} f_N(x).$$

This is justified by the following result:

Theorem 1. Consider equation (1) and assume that

(9)
(a) $g(0) = h(0) = 0,$

(b) $0 < a,\ b < 1$

(c) $g(x)$ and $h(x)$ are continuous and monotone increasing in $[0, x_0]$.

(d) $\sum_{n=0}^{\infty} g(c^n x) < \infty, \quad \sum_{n=0}^{\infty} h(c^n x) < \infty,$

where $c = \text{Max}(a,b)$.

Under these conditions there is a unique solution which is continuous in $[0, x_0]$ and possesses the value 0 at 0.

If

$$(10)\qquad f_1 = \operatorname*{Max}_{0 \leqq y \leqq x} \left[g(y) + h(x-y) \right]$$

$$f_{n+1} = \operatorname*{Max}_{0 \leqq y \leqq x} \left[g(y) + h(x-y) + f_n(ay + b(x-y)) \right],$$

we have

$$(11)\qquad f(x) = \lim_{n \to \infty} f_n(x).$$

For the proof of this and the five results stated below, we refer to [2] and [6].

§5. Approximation Techniques— Successive Approximations-I.

Let us write our functional equation in the form

$$(12)\qquad f = T(f, P),$$

where f represents the unknown function, T represents the transformation $\operatorname*{Max}_{0 \leqq y \leqq x} \left[g(y) + h(x-y) + f(ay + b(x-y)) \right]$, and P represents the set of known parameters, the functions $g(x)$ and $h(x)$ and the constants a and b, that appear in T.

In theory there is only one method to be used in approximating the solution of a functional equation, namely the technique of solving an approximate functional equation. It is in the choice of these approximate equations that practice varies.

The method of successive approximations in its usual guise relies upon solving the following system of equations:

$$(13) \qquad f_{n+1} = T(f_n, P),$$

where f_1 is a guess at the solution. In more refined applications, (13) is replaced by

$$(14) \qquad f_{n+1} - R(f_{n+1}) = T(f_n, P) - R(f_n)$$

where R is a transformation so chosen as to force f_n to possess certain desired properties or to increase the rapidity of convergence.

A simple way to proceed is to mimic the physical process and take

$$(15) \qquad f_1 = \underset{0 \leq y \leq x}{\mathrm{Max}} \left[g(y) + h(x-y) \right],$$

and

$$(16) \qquad f_{n+1} = \underset{0 \leq y \leq x}{\mathrm{Max}} \left[g(y) + h(x-y) + f_n(ay + b(x-y)) \right].$$

The computations are quite easy to perform and possess the merit of furnishing useful information at the same time. However, it is not to be expected that the convergence will be very rapid initially. Consequently, we shall investigate some other procedures.

§6. Approximation Techniques—Simplified Equation.

In place of the above approximation, we may approximate by replacing the equation $f = T(f,P)$ by

$$(17) \qquad f = T(f,P')$$

where P' represents a different set of parameters, one which permits a solution in toto, or which yields a stronger hold on the solution.

Thus, for example, in our equation

$$(18) \qquad f(x) = \operatorname*{Max}_{0 \leq y \leq x} \left[g(y) + h(x-y) + f(ay + b(x-y)) \right],$$

we make the further assumption that g and h are convex functions. The following result then holds:

Theorem 2. If g and h are convex functions and the conditions of Theorem 1 hold, an optimal policy consists of choosing $y = 0$ or x.

Although this simplifies the finding of the solution, it is still not easy to find an explicit solution.

If we wish to obtain an explicit approximate solution, we can make the further approximation that

$$(19) \qquad g = ax^b, \quad h = cx^d,$$

This corresponds to the approximation of $\log g(x)$ by $\alpha_1 + \beta_1 \log x$ or $\log g(e^x)$ by $\alpha_1 + \beta_1 x$.

In the case where g and h have the simple forms given in (19) above, we have the following result:

Theorem 3. The solution of

$$(20) \qquad f(x) = \operatorname{Max}\left[cx^d + f(ax),\quad ex^f + f(bx)\right]$$

subject to

$$(21) \qquad \text{(a)} \quad 0 < a,\ b < 1, \quad c,d > 0$$

$$\text{(b)} \quad 0 < d < f$$

is given by

$$(22) \qquad \begin{aligned} f(x) &= cx^d + f(ax), && 0 \leq x \leq x_0 \\ &= ex^f + f(bx), && x_0 \leq x \end{aligned}$$

where

$$(23) \qquad x_0 = \left[\frac{\dfrac{c}{1-a^f}}{\dfrac{e}{1-b^d}}\right]^{1/(f-d)}$$

In the general case where g and h are convex and we know that y = 0 or x at each stage, partial results similar to the above can be found. It would be interesting to know under what further assumption in addition to convexity one has a solution similar to (22).

If g and h are not convex, but concave, we may use the following result to obtain a solution:

<u>Theorem</u> 4. <u>Let</u>

$$\text{(24)}\qquad \begin{aligned} &\text{(a)}\quad g(0) = h(0) = 0, \\ &\text{(b)}\quad g'(x),\ h'(x) \geqq 0 \quad \underline{\text{for}}\ x \geqq 0, \\ &\text{(c)}\quad g''(x),\ h''(x) \leqq 0 \quad \underline{\text{for}}\ x \geqq 0, \end{aligned}$$

<u>and consider the sequence of approximations to</u> f <u>defined by</u>

$$\text{(25)}\qquad \begin{aligned} f_0(x) &= \operatorname*{Max}_{0 \leqq y \leqq x} \left[g(y) + h(x-y) \right], \\ f_{n+1}(x) &= \operatorname*{Max}_{0 \leqq y \leqq x} \left[g(y) + h(x-y) + f_n\left[ay + b(x-y)\right] \right], \\ & \qquad\qquad n = 0, 1, 2, \cdots . \end{aligned}$$

<u>For each</u> n, <u>there is a unique</u> $y_n = y_n(x)$ <u>that yields the maximum. If</u> $b \leqq a$, <u>we have</u> $y_1 \leqq y_2 \leqq y_3 \cdots$, <u>and the reverse inequalities for</u> $b \geqq a$. <u>In particular, if</u> $y_n(x) = x$ <u>for some</u> n <u>in the case</u> $b \leqq a$, <u>then</u> $y_m(x) = x$ <u>for</u> $m \geqq n$, <u>and the solution of the original equation in</u> (16) <u>will be furnished by</u> $y = x$.

Let us note finally that if an interior maximum exists we must have simultaneously

$$\text{(26)}\qquad \begin{aligned} & g'(y) - h'(x-y) + (a-b) f'(ay + b(x-y)) = 0 \\ & f'(x) = h'(x-y) + b f'(ay + b(x-y)) . \end{aligned}$$

These equations may be solved explicitly for y and f if g and h are quadratic.

§7. Approximation Technique—Approximation in Strategy Space.

Up to now we have been discussing conventional approximation techniques, common to the functional equations that arise in mathematical physics. Let us now discuss a technique that is particularly suited to dynamic programming.

In following the above approach we established an equivalence between the space of all allowable allocations, a strategy space, and the function space of all conceivable solutions of our functional equation. An optimal policy yields, by direct computation, the solution of the functional equation, and conversely, the functional equation, through its determination of $y(x)$, yields an optimal sequence of allocations.

It follows then, that we have a duality between the strategy space and the function space with the prerogative of attacking the problem on the grounds of our own choosing.

This immediately furnishes us with a new, powerful technique for finding approximate solutions. In place of approximating in function space we may approximate in strategy space. It is in this way that we may most efficiently exploit the insight and intuition gained from experience.

For example, we might argue that the unit cost is the determining factor and set $y = 0$ whenever

$$\frac{g(x)}{(1-b)x} > \frac{h(x)}{(1-a)x}, \tag{27}$$

and $y = x$ otherwise.

Using this policy we compute a function $f_1(x)$. This is now used as a first approximation.

The great advantage of this technique lies in the fact that it ensures monotone convergence. We know automatically that the next approximation will yield a superior policy. To demonstrate this, let $f_1(x)$ be generated by a rule which furnishes y given x. Then

(28) $$f_1(x) = g(y) + h(x-y) + f_1(ay + b(x-y)).$$

It follows that if $f_2(x)$ is determined by

(29) $$f_2(x) = \max_{0 \leq y \leq x} \left[g(y) + h(x-y) + f_1(ay + b(x-y)) \right],$$

we have $f_2 \geq f_1$ with equality only if f_1 is the actual solution. Having established that $f_2 \geq f_1$ it is immediate that f_3 as determined by

(30) $$f_3(x) = \max_{0 \leq y \leq x} \left[g(y) + h(x-y) + f_2(ay + b(x-y)) \right]$$

is greater than or equal to f_2, and inductively that $f_{n+1} \geq f_n$ for $n=1,2,\cdots$.

The whole point of solving a simple model of a decision problem is not so much that it furnishes an approximate function, but rather that it furnishes an approximate policy, which is now used to furnish an approximate function for a more complicated and realistic problem.

§8. Some Generalizations.

Let us consider several immediate generalizations. We may first of all consider the case where the return and the cost are both functions of the stage. The resultant functional equations then have the form

$$(31) \qquad f_k(x) = \max_{0 \leq y \leq x} \left[a_k(x,y) + f_{k+1}\{b_k(x,y)\} \right].$$

A more interesting generalization is that where the return is not determined, but subject to a probability distribution.

Thus, as an illustration, let us assume that if the initial allocation is y there is a probability p_1 that the return is $g_1(y) + h_1(x-y)$ and that the quantity remaining is $a_1y + b_1(x-y)$, and a probability $p_2 = 1 - p_1$ that the return is $g_2(y) + h_2(x-y)$ and the quantity left is $a_2y + b_2(x-y)$.

Since the return is now a stochastic quantity, it is no longer possible to speak of maximizing the return, but rather to speak of maximizing the average value of some function of this return. The simplest measure is the expected return. Let

(32) $f(x)$ = expected total return obtained using an optimal policy.

Then, as above, we obtain the functional equation

$$(33) \qquad f(x) = \max_{0 \leq y \leq x} \Big[p_1 \left\{ g_1(y) + h_1(x-y) + f(a_1y + b_1(x-y)) \right\} + p_2 \left\{ g_2(y) + h_2(x-y) + f(a_2y + b_2(x-y)) \right\} \Big]$$

Results analogous to those described in the preceding sections hold for this and the still more general form

$$(34) \qquad f(x) = \max_{0 \le y \le x} \left[\int_0^1 \left[a(x,y,z) + f(b(x,y,z)) \right] dG(z,y) \right],$$

where the distribution of outcomes depends upon the outcome.

Functional equations of similar type occur in the work on the optimal inventory problem of Arrow, Harris, and Marschak [1], and Dvoretzky, Kiefer, and Wolfowitz [9].

§9. A Particular Example.

In the previous sections we have discussed approximation techniques which are particularly applicable when g and h are either concave or convex. Since, in applications, curves with points of inflection are of frequent occurrence, it is of some interest to see what occurs when g and h have neither of the simple forms.

A particularly simple pair of easily computable functions possessing points of inflection are

$$(35) \qquad g(y) = re^{-c/y}, \qquad h(y) = e^{-d/y}.$$

The equation now has the form

$$(36) \qquad f(x) = \max_{0 \le y \le x} \left[re^{-c/y} + e^{-d/(x-y)} + f(ay + b(x-y)) \right],$$

with

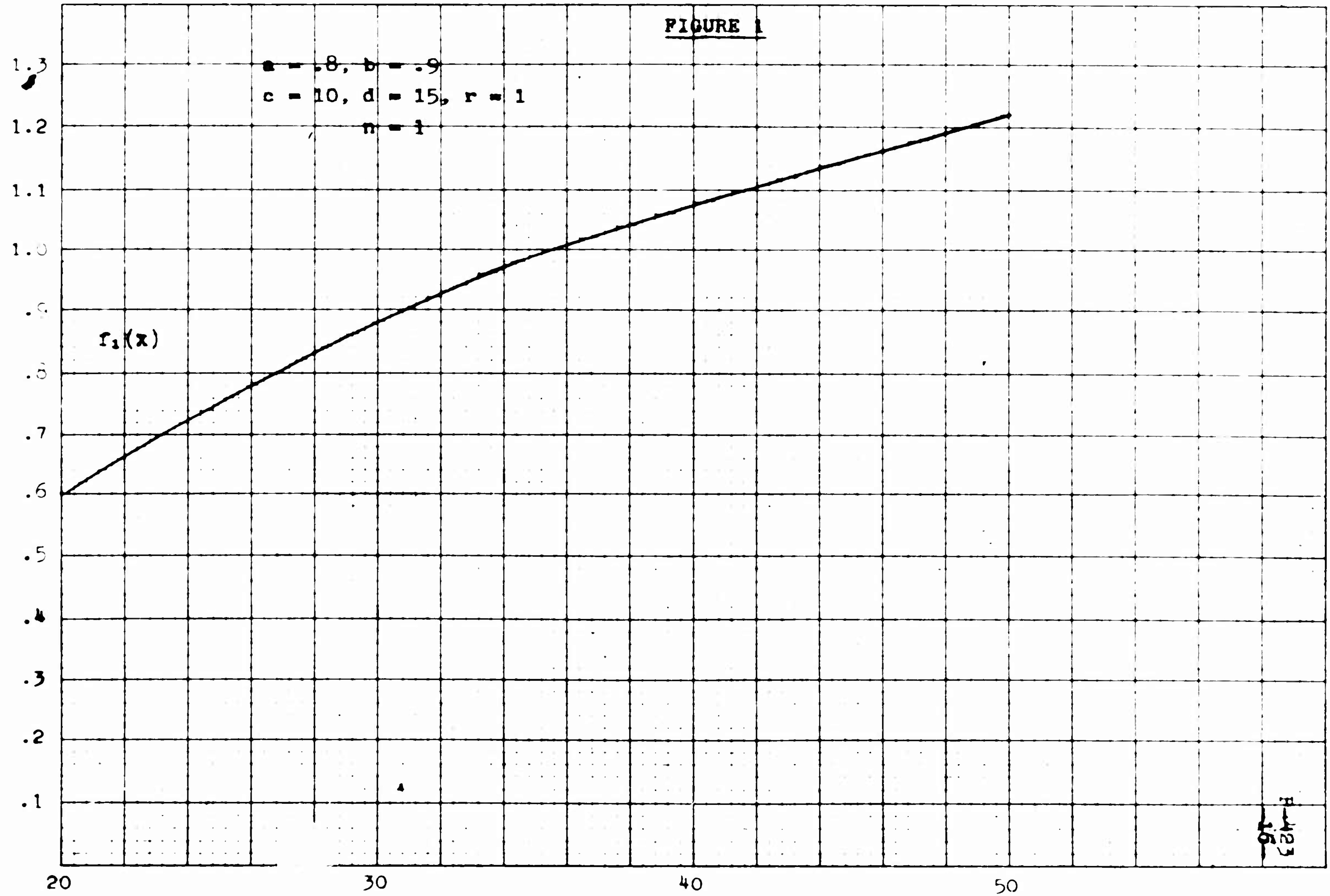
$$(37) \qquad f_1(x) = \underset{0 \le y \le x}{\text{Max}} \left[re^{-c/y} + e^{-d/(x-y)} \right],$$

$$f_2(x) = \underset{0 \le y \le x}{\text{Max}} \left[re^{-c/y} + e^{-d/(x-y)} + f_1(ay + b(x-y)) \right].$$

These functions were computed for various sets of values; the curves for $a = .8$, $b = .9$, $r = 1$, $c = 10$, $d = 15$ are appended. What is striking is that although the graphs of $f(x)$, $f_1(x)$, $f_2(x)$ are quite smooth, the graph of $y = y(x)$, the maximizing choice of y, is quite disjointed. It is probably true, although we have not verified it, that there exists an approximate strategy which is slowly varying in x and yields almost as large a total return as the exact strategy. This property is typical of many problems of this type.

bjc

FIGURE 1



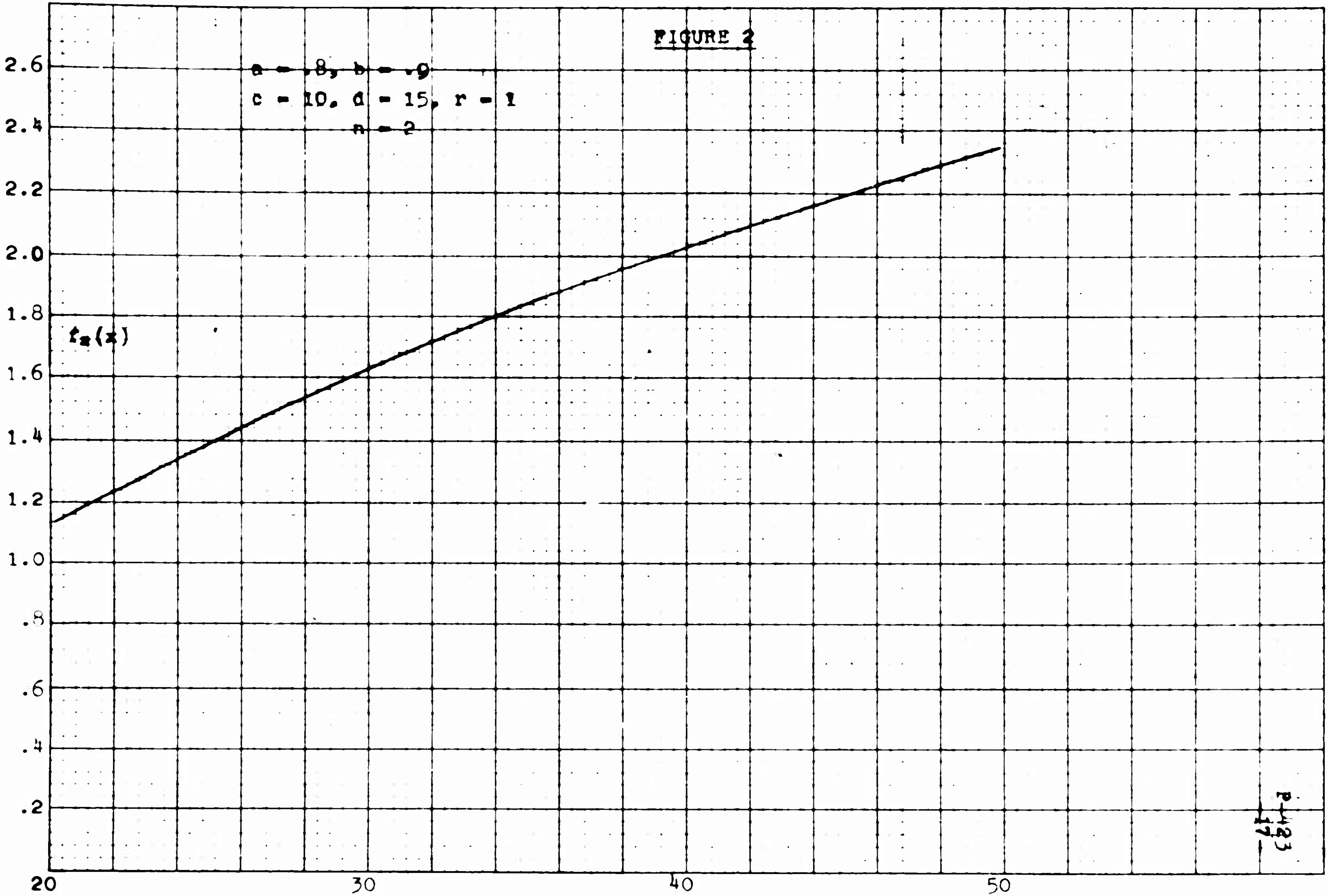


FIGURE 2

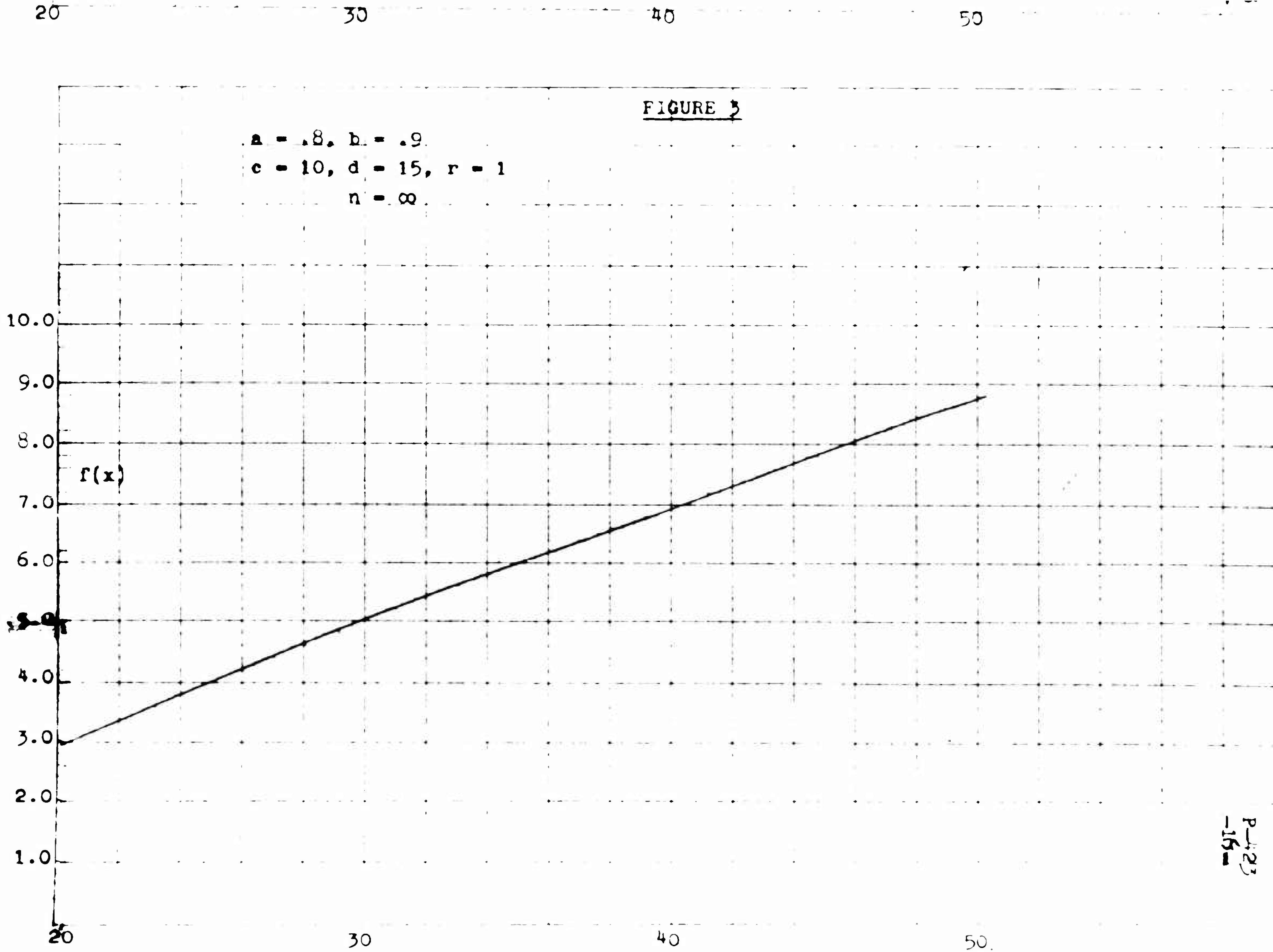
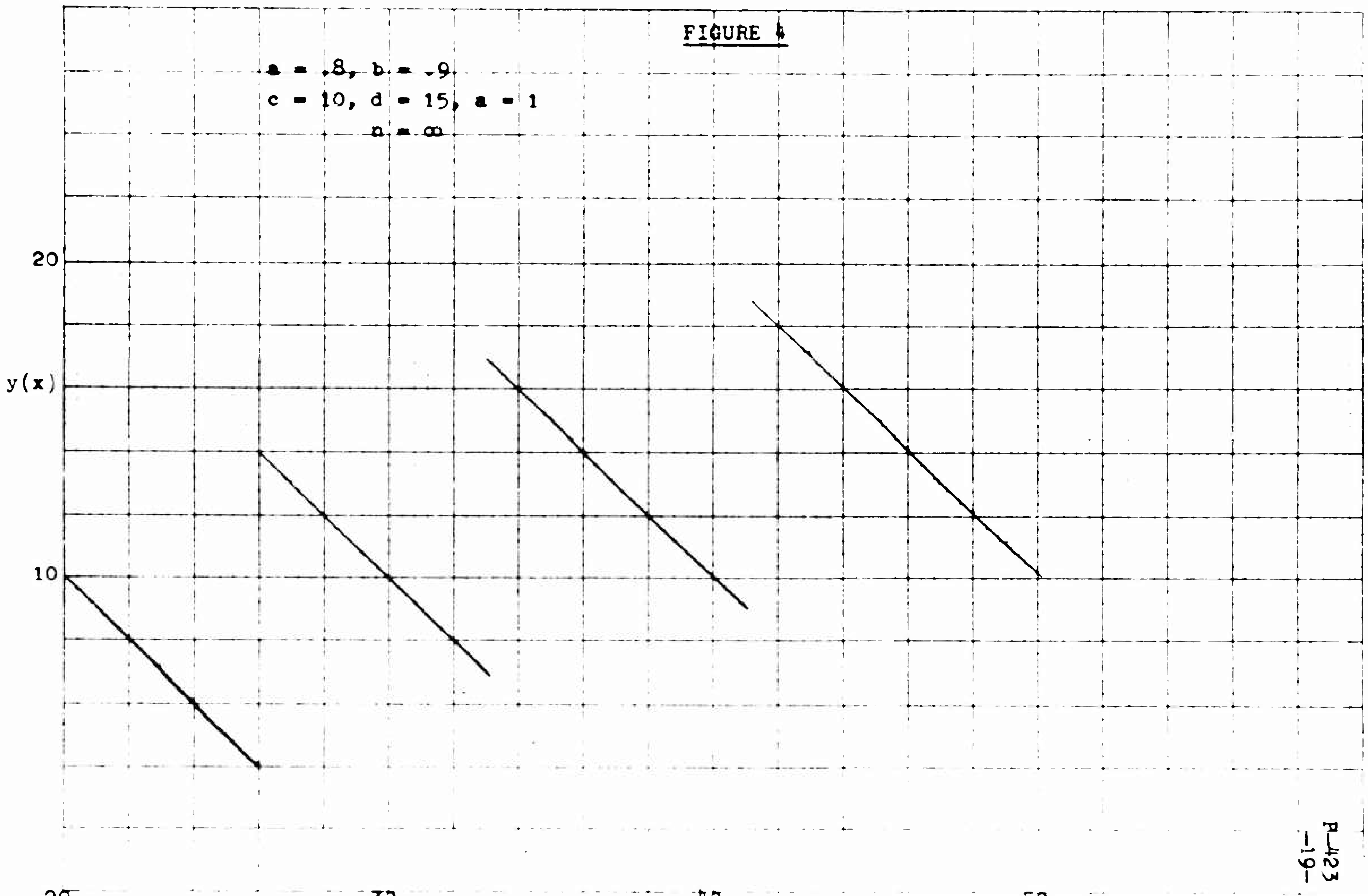


FIGURE 3

FIGURE 4



BIBLIOGRAPHY

1. Arrow, K. J., Harris, T. E., and Marschak, J. "Optimal Inventory Policy," Cowles Commission Paper No. 44, 1951.

2. Bellman, R. An Introduction to the Theory of Dynamic Programming. The RAND Corporation Report No. 245, 1953.

3. ———. "The Theory of Dynamic Programming," Proc. Nat. Acad. Sci., Vol. 38, No. 8, August 1952, pp. 716–719.

4. ———. "Some Functional Equations in the Theory of Dynamic Programming," Proc. Nat. Acad. Sci. (to appear).

5. ———. "On Bottleneck Problems and Dynamic Programming," Proc. Nat. Acad. Sci. (to appear).

6. ———. "The Theory of Dynamic Programming," Econometrica, (to appear).

7. Bellman, R., Glicksberg, I., and Gross, O. "Some Variational Problems Occurring in the Theory of Dynamic Programming," Proc. Nat. Acad. Sci., Vol. 39, No. 4, April 1953, pp. 298–301.

8. ———. "On Some Variational Problems Occurring in the Theory of Dynamic Programming," Rendiconti del Circolo Matematico di Palermo, (to appear).

9. Dvoretzky, A. J., Kiefer, J., and Wolfowitz, J. "The Inventory Problem—I: Case of Known Distribution of Demand," and "The Inventory Problem—II: Case of Unknown Distributions of Demand," Econometrica, Vol. 20, No. 2, April 1952, pp. 187–222.